Optimal Piecewise Affine Approximations of Nonlinear Functions Obtained from Measurements

Alexander Szűcs [*,1] **Michal Kvasnica** [*], and **Miroslav Fikar** [*]

[*] *Institute of Automation, Information Engineering and Mathematics, Slovak University of Technology, 812 37 Bratislava, Slovakia*

Abstract: The paper describes a two-stage procedure for obtaining piecewise affine approximations of static nonlinearities obtained from measured data. In the first step we search for a suitable function which fits the data while minimizing the fitting error. Subsequently we show how to approximate, in an optimal fashion, the nonlinear fitting function by a piecewise affine function of pre-specified complexity. We illustrate that approximation of arbitrary nonlinear functions boils down to a series of one-dimensional approximations, rendering the procedure efficient from a computational point of view.

Keywords: hybrid systems, approximation, static nonlinearities

1. INTRODUCTION

Hybrid systems (Branicky, 1995) nowadays represent a proven mathematical framework for modeling of complex systems which include interconnection between continuous dynamics and discrete logic. Applications include, but are not limited to, power electronics (Papafotiou et al., 2007) or automotive (Corona and De Schutter, 2008) systems. Popularity of hybrid systems is mainly due to the fact that they provide an accurate description of plant's dynamics while simultaneously allowing for simple analysis and control synthesis (Bemporad and Morari, 1999).

One of frequently employed modeling frameworks for hybrid systems are Piecewise Affine systems (PWA) (Sontag, 1981), where the space of variables is partitioned into a finite number of non-overlapping regions, each of them associated with a linear (or affine) expression. The main advantage of PWA models is their ability to approximate arbitrary nonlinearities. Therefore, significant research activity has been devoted to developing techniques for construction of PWA approximations of nonlinearities. Two main directions can be found in the literature.

If the analytical form of the nonlinearity is known, one option is to construct the PWA approximation *manually* using the HYSDEL language (Torrisi and Bemporad, 2004). As an alternative, in our previous work (Kvasnica et al., 2011) we have shown how to build optimal PWA approximations *automatically* by solving nonlinear optimization problems.

In many practical instances, though, the nonlinearity to be approximated is not available in its analytic form. Alternatively, the analytic form could be known, but its numerical parameters are not. In such cases one usually resorts to obtaining PWA approximations directly from measurements where the nonlinearity characteristics has to be extracted from input-output data. Such approaches are usually at the core of most PWA identification techniques, see e.g. Ferrari-Trecate et al. (2001); Roll et al. (2004); Paoletti et al. (2007); Gegúndez et al. (2008); Ohlsson et al. (2010). These approaches, however, have two crucial downsides. First, they are typically time consuming since they rely on solving high-dimensional optimization problems. Second, in order to obtain a well-defined PWA approximation, the procedure has to determine regions of the PWA model which do not overlap and whose union covers the whole space of parameters of interest, without leaving "holes" where the model would be undefined. The latter requirement is difficult to guarantee, further increasing complexity of these schemes.

To overcome these difficulties, in this paper we propose to extend our previous work (Kvasnica et al., 2011) by using a two-stage optimization-based technique to derive PWA approximations of *static* nonlinearities obtained from measured data. The first part of the procedure is focused on finding the best fit of measured data by a pre-specified set of basis functions. A similar idea was suggested in Kozák and Števek (2011) where the authors employed neural networks to find the fitting function of a particular structure. In this paper we advocate to find the fit by solving standard optimization problems. Moreover, we also show how to find the fitting function of minimal complexity by solving a binary optimization problem. The result of this stage is an analytical formula of the fitting function which is used as an input to the second step. There, following our previous work we show how to approximate an arbitrary nonlinear function by a PWA model such that the approximation error is minimized. Moreover, we show that, regardless of dimensionality of the function to be approximated, the approximation procedure always boils down to a series of one-dimensional approximations, keeping complexity of the presented approach on acceptable level. In addition to Kvasnica et al. (2011) we also show how to derive PWA approximations of nonlinear functions

[1] Corresponding author, e-mail:alexander.szucs@stuba.sk

which are not separable. This significantly extends scope of the proposed work.

2. PROBLEM STATEMENT

In this paper we aim at addressing the following problem. We are given T samples of input data $z_i \in \mathcal{Z} \subset \mathbb{R}^{n_z}$ from some closed and bounded set $\mathcal{Z}$, and the corresponding measurements $y_i \in \mathbb{R}$, $i = 1, \ldots, T$. We want to fit the data with a PWA function $\tilde{f} : \mathbb{R}^{n_z} \to \mathbb{R}$ with N regions

$$\tilde{f}(z) = \begin{cases} a_1^T z + c_1 & \text{if } z \in \mathcal{R}_1, \\ & \vdots \\ a_N^T z + c_N & \text{if } z \in \mathcal{R}_N, \end{cases} \tag{1}$$

which satisfies two design requirements:

R1: $\tilde{f}$ is well-posed (Bemporad and Morari, 1999) on $\mathcal{Z}$, i.e it satisfies $\mathrm{int}(\mathcal{R}_i) \cap \mathrm{int}(\mathcal{R}_j) = \emptyset$, $\forall i \neq j$ and $\cup_j \mathcal{R}_j = \mathcal{Z}$, $j = 1, \ldots, N$.

R2: $\tilde{f}$ is a good fit which achieves a low fitting error $e_{\text{fit}} = \sum_{i=1}^{T} (y_i - \tilde{f}(z_i))^2$.

Solving this problem (i.e. determining regions $\mathcal{R}_j \subseteq \mathbb{R}^{n_z}$ and parameters $a_j \in \mathbb{R}^{n_z}$, $c_j \in \mathbb{R}$, $j = 1, \ldots, N$), however, is not trivial (Kvasnica et al., 2011) if the input samples z_i are vectors, i.e. when $n_z > 1$. The difficult part is how to divide the domain $\mathcal{Z}$ into non-overlapping regions $\mathcal{R}_j$ without creating "holes", i.e. guaranteeing that the union $\cup_j \mathcal{R}_j$ completely covers $\mathcal{Z}$ if dimension$(\mathcal{Z}) > 1$.

To overcome this difficulty, we propose to split the search for the PWA function $\tilde{f}$ into two steps. In the first stage we fit the input data, represented by the (z_i, y_i) pairs, with a nonlinear function $y = f(z)$, where $f : \mathbb{R}^{n_z} \to \mathbb{R}$ is given by

$$f(z) = \alpha_1 f_1(z) + \cdots + \alpha_n f_n(z), \tag{2}$$

where $f_i : \mathbb{R}^{n_z} \to \mathbb{R}$ are pre-specified basis functions, and $\alpha_i \in \mathbb{R}$ are scalar coefficients.

Problem 2.1. Given are T samples of input-output data (z_i, y_i), $i = 1, \ldots, T$. Determine coefficients α_i of f in (2) such that the fitting error

$$e_{\text{fit}} = \sum_{i=1}^{T} (y_i - f(z_i))^2 \tag{3}$$

is minimized.

Once the analytical form of the fitting function f is available, in the second step we search for its optimal PWA approximation:

Problem 2.2. Given is a nonlinear function $f : \mathbb{R}^{n_z} \to \mathbb{R}$ and its domain $\mathcal{Z} \subset \mathbb{R}^{n_z}$. Find a well-posed PWA approximation $\tilde{f}$ of *pre-specified* complexity N as in (1) such that the approximation error

$$e_{\text{aprx}} = \int_{\mathcal{Z}} (f(z) - \tilde{f}(z))^2 \mathrm{d}z \tag{4}$$

is minimized.

3. FUNCTION FITTING

To solve Problem 2.1 we need to determine coefficients $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, n$ which parametrize f and provide an optimal fit in the sense of minimizing (3). Needless to say, selection of the basis functions is crucial in obtaining a good fit. In many situations the basis is chosen by hand, employing prior knowledge about the analytical form of the nonlinearity from which the input-output data originated. One such an example will be provided in Section 5.

If this prior information is not available, one can resort to a rather broad selection of basis functions (Boyd and Vandenberghe, 2004, Ch. 6). One common subspace of functions on $\mathbb{R}$ consists of polynomials of degree less than n. The simplest basis consists of the powers, i.e. $f_i(z) = z^{i-1}, i = 1, \ldots, n$. We can also consider polynomials on $\mathbb{R}^{n_z}$, with a maximum total degree n

$$f_i(z) = \sum_{i_1 + \cdots + i_n \leq n} z_1^{i_1} \cdots z_{n_z}^{i_n}, \tag{5}$$

or a maximum degree for each variable.

Regardless of the choice of the basis functions, it is important to notice that f as in (2) is linear in the unknown coefficients $\alpha_1, \ldots, \alpha_n$. Therefore Problem 2.1 can be easily solved by directly minimizing (3) e.g. by taking the derivative equal to zero.

Problem 2.1 can be further extended as to find f which consists of the least possible number of basis functions, i.e. by minimizing the cardinality of the vector of parameters $\alpha = [\alpha_1, \ldots, \alpha_n]$ in (2). A simple heuristic approach would be to minimize the 1-norm of α (Boyd and Vandenberghe, 2004):

$$\min \sum_{i=1}^{T} (y_i - f(z_i))^2 + \gamma \|\alpha\|_1, \tag{6}$$

which can be cast as a constrained quadratic program. The tuning parameter $\gamma > 0$ here acts as a regularization coefficient.

A more rigorous approach is to directly minimize the number of non-zero components of α. This can be achieved by introducing a set of binary indicators $\delta_j \in \{0, 1\}$, $j = 1, \ldots, n$ which fulfill

$$(\alpha_j \neq 0) \Rightarrow (\delta_j = 1). \tag{7}$$

By employing the big-M technique (Williams, 1993; Bemporad and Morari, 1999) we can rewrite (7) into a set of inequalities which are linear in δ_j and α_j:

$$-M\delta_j \leq \alpha_j \leq M\delta_j, \tag{8}$$

where M is a sufficiently large number. It is then easy to verify that minimization of the number of nonzero components amounts to minimizing the sum of corresponding binary indicators, i.e.

$$\min \sum_{i=1}^{T} (y_i - f(z_i))^2 + \gamma \sum_{j=1}^{n} \delta_j \tag{9a}$$

$$\text{s.t. } -M\delta_j \leq \alpha_j \leq M\delta_j, \; j = 1, \ldots, n, \tag{9b}$$

which provides a good fit of minimal cardinality. Problem (9) is a mixed-integer quadratic program which can be solved to global optimality using state-of-the-art solvers (Löfberg, 2004; ILOG, Inc., 2003). Complexity of (9) is primarily determined by the number of binary variables, i.e. by the number n of basis functions considered in (2).

4. OPTIMAL PWA APPROXIMATION

In this section we show how to solve Problem 2.2 provided that the analytical form of f is known. Results of this section therefore cover the scenario where f was obtained by the fitting procedure of Section 3, but also apply to situations where the nonlinearity stems from an a-priori known analytical relation.

We distinguish between three cases. The first one, described in Section 4.1, covers approximation of one-dimensional functions where $f : \mathbb{R} \to \mathbb{R}$ provided that the domain of f is connected and closed. Then, in Section 4.2 we show how to extend the procedure to approximation of multi-variable functions given as products of functions of single variable. Finally, in Section 4.3 we illustrate how to solve Problem 2.2 where f is an arbitrarily complex function, not satisfying any special properties.

4.1 Functions in One Variable

First, we consider the one-dimensional case, i.e. approximation of a nonlinear function $f : \mathbb{R} \to \mathbb{R}$, with domain $\mathcal{Z} \subset \mathbb{R}$, by a PWA function $\tilde{f}$ as in (1). Since $\mathcal{Z}$ is assumed to be connected and closed, it is a line segment $[\underline{z}, \overline{z}]$. Regions $\mathcal{R}_i$ define the partition of such a line into N non-overlapping parts, i.e. $\mathcal{R}_1 = [\underline{z},\ r_1]$, $\mathcal{R}_2 = [r_1,\ r_2]$, $\ldots$,$\mathcal{R}_{N-1} = [r_{N-2},\ r_{N-1}]$, $\mathcal{R}_N = [r_{N-1},\ \overline{z}]$. Such a subdivision trivially satisfies $\mathrm{int}(\mathcal{R}_i) \cap \mathrm{int}(\mathcal{R}_j) = \emptyset$, $\forall i \neq j$ and $\cup_j \mathcal{R}_j = \mathcal{Z}$, $j = 1, \ldots, N$.

Solving Problem 2.2 then becomes to find slopes a_i, offsets c_i and breakpoints r_i such that the approximation error is minimized, i.e.

$$\min_{a_i, c_i, r_i} \sum_{i=1}^{N} \Big(\int_{r_{i-1}}^{r_i} \big(f(z) - (a_i z + c_i)\big)^2 \, dz \Big) \tag{10a}$$

$$\text{s.t. } \underline{z} \leq r_1 \leq \cdots \leq r_{N-1} \leq \overline{z}, \tag{10b}$$

$$a_i r_i + c_i = a_{i+1} r_i + c_{i+1}, \tag{10c}$$

with $r_0 = \underline{z}$ and $r_N = \overline{z}$.

For simple functions f, the integral in (10a) can be expressed in an analytical form in unknowns a_i, c_i, r_i, along with the corresponding gradients. For more complex expressions, the integrals can be evaluated numerically, e.g. by using the trapezoidal rule. In either case, problem (10) can be formulated as a nonlinear optimization problem (NLP) and solved to a local optimality e.g. by using the `fmincon` solver of MATLAB. Alternatively, one can use global optimization methods (Adjiman et al., 1996; Papamichail and Adjiman, 2004; Chachuat et al., 2006), which guarantee that an ϵ-neighborhood of the global optimum can be found.

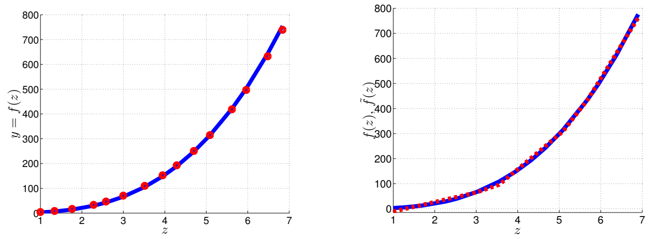
Remark 4.1. Constraint (10c) guarantees that the PWA approximation is continuous. If a discontinuous approximation is desired, the constraint can be omitted or modified to explicitly account for discontinuity.

Example 4.1. Consider a set of input-output data $(z_i,\ y_i)$ shown in Figure 1(a). To fit these data with a function f as in (2) we have selected a set of polynomial basis functions $f_i(z) = z^{i-1}$ with a maximum degree $n = 3$ such that

$$f(z) = \alpha_1 + \alpha_2 z + \alpha_3 z^2 + \alpha_4 z^3. \tag{11}$$

By minimizing e_{fit} in (3) we have obtained the optimal fit $f(z) = 0.22 + 0.08z + 0.42z^2 + 2.29z^3$, shown in Figure 1(a). The simplest possible fit could also be obtained by solving (9), which resulted in $f(z) = 2.40z^3$. Subsequently, we have obtained the PWA approximation $\tilde{f}$ by solving the NLP (10) by considering $N = 3$ regions of $\tilde{f}$. The resulting optimal PWA approximation of this complexity is shown in Figure 1(b) and is given by

$$\tilde{f}(z) = \begin{cases} 38.1041z - 49.1898 & \text{if } 1 \leq z \leq 3.4547 \\ 138.0073z - 394.329 & \text{if } 3.4547 \leq z \leq 5.3452 \\ 267.9618z - 1088.9654 & \text{if } 5.3452 \leq z \leq 7 \end{cases}$$



(a) Input-output data $(z_i,\ y_i)$ and optimal fit with a polynomial basis of maximum degree 3.

(b) Graph of $f(z)$ (solid line) and its optimal PWA approximation $\tilde{f}(z)$ with 3 regions (red dashed line).

Fig. 1. One-dimensional fit from Example 4.1.

4.2 Multivariable Separable Functions

Next we show how to derive optimal PWA approximations of multivariable functions $f(z_1, \ldots, z_{n_z}) : \mathbb{R}^{n_z} \to \mathbb{R}$ with domain $\mathcal{Z} \subset \mathbb{R}^{n_z}$, provided that the analytical form of f satisfies the following condition:

Assumption 4.1. The multivariable nonlinear function f can be represented as a sum of products of functions in single variables, i.e. $f(z_1, \ldots, z_{n_z}) = \sum_{i=1}^{n_z} \alpha_i \Big(\prod_{j=p_i}^{q_i} f_j(z_j) \Big)$. Here, α_i are scalar coefficients and $f_j : \mathbb{R} \to \mathbb{R}$ are scalar-valued (possibly nonlinear) basis functions.

One special case of Assumption 4.1 are so-called *separable* functions (Williams, 1993) where f can be expressed as a sum of functions of a single variable, i.e. $f(z_1, \ldots, z_n) = f_1(z_1) + \cdots + f_n(z_n)$. If f is readily separable (e.g. when $f(z_1, z_2) = e^{z_1} + \sin(z_2)$), its optimal PWA approximation can be easily obtained by applying the 1D scenario of Section 4.1 to the individual components of the function, i.e. $\tilde{f}(z_1, \ldots, z_{n_z}) = \tilde{f}_1(z_1) + \cdots + \tilde{f}_{n_z}(z_{n_z})$. The total number of regions over which the PWA approximation $\tilde{f}(\cdot)$ is defined is hence given by $\sum_{j=1}^{n} N_j$, where N_j is the pre-specified complexity of the j-th approximation $\tilde{f}_j(z_j)$.

If f is not separable, but satisfies Assumption 4.1, it can be converted into the separable form by applying a simple change of variables, elaborated in more details e.g. in Williams (1993). To illustrate the procedure, consider a non-separable function $f(z_1, z_2) = z_1 z_2$ with domain $\mathcal{Z} := [\underline{z}_1,\ \overline{z}_1] \times [\underline{z}_2,\ \overline{z}_2]$. Define two new variables

$$y_1 = (z_1 + z_2), \quad y_2 = (z_1 - z_2). \tag{12}$$

Then it is easy to verify that $1/4(y_1^2 - y_2^2) = z_1 z_2$. The coordinate transformation therefore transforms the original function into a separable form, where both terms (y_1^2 and y_2^2) are now functions of a single variable. The procedure of Section 4.1 can thus be applied to compute PWA approximations of $f_{y_1}(y_1) := y_1^2$ and $f_{y_2}(y_2) := y_2^2$, where the function arguments relate to z_1 and z_2 via (12). Important to notice is that $f_{y_1}(\cdot)$ and $f_{y_2}(\cdot)$ have different domains, therefore their PWA approximations $\tilde{f}_{y_1}(y_1) \approx y_1^2$ and $\tilde{f}_{y_2}(y_2) \approx y_2^2$ will, in general, be different. Specifically, the domain of $f_{y_1}(\cdot)$ is $[\underline{y}_1,\ \overline{y}_1]$ with $\underline{y}_1 = \min\{z_1 + z_2 \mid \underline{z}_1 \le z_1 \le \overline{z}_1,\ \underline{z}_2 \le z_2 \le \overline{z}_2\}$ and $\overline{y}_1 = \max\{z_1+z_2 \mid \underline{z}_1 \le z_1 \le \overline{z}_1,\ \underline{z}_2 \le z_2 \le \overline{z}_2\}$. Similarly, the domain of $f_{y_2}(\cdot)$ is $[\underline{y}_2,\ \overline{y}_2]$, whose boundaries can be computed by respectively minimizing and maximizing $z_1 - z_2$ subject to the constraint $[z_1,\ z_2]^T \in \mathcal{Z}$. The overall PWA approximation $\tilde{f}(z_1, z_2) \approx z_1 z_2$ then becomes

$$\tilde{f}(z_1, z_2) = {}^1\!/\!_4(\tilde{f}_{y_1}(z_1 + z_2) - \tilde{f}_{y_2}(z_1 - z_2)). \tag{13}$$

The value of $\tilde{f}(z_1, z_2)$ for any points z_1, z_2 is obtained by subtracting the value of the PWA function $\tilde{f}_{y_2}(\cdot)$ evaluated at the point $z_1 - z_2$ from the function value of $\tilde{f}_{y_1}(\cdot)$ evaluated at $z_1 + z_2$, followed by a linear scaling.

The procedure naturally extends to multivariable functions represented by the product of two nonlinear functions of a single variable, i.e. $f(z_1, z_2) = f_1(z_1)f_2(z_2)$. Here, the transformation (12) becomes

$$y_1 = f_1(z_1) + f_2(z_2), \quad y_2 = f_1(z_1) - f_2(z_2). \tag{14}$$

Therefore, ${}^1\!/\!_4(y_1^2 - y_2^2) = f(z_1, z_2)$ still holds. Let $f_{y_1}(y_1) := y_1^2$ and $f_{y_2}(y_2) := y_2^2$. The domain of $f_{y_1}(\cdot)$ is $[\underline{y}_1,\ \overline{y}_1]$ and $\mathrm{dom} f_{y_2}(\cdot) = [\underline{y}_2,\ \overline{y}_2]$ with

$$\underline{y}_1 = \min\{f_1(z_1) + f_2(z_2) \mid [z_1,\ z_2]^T \in \mathcal{Z}\}, \tag{15a}$$
$$\overline{y}_1 = \max\{f_1(z_1) + f_2(z_2) \mid [z_1,\ z_2]^T \in \mathcal{Z}\}, \tag{15b}$$
$$\underline{y}_2 = \min\{f_1(z_1) - f_2(z_2) \mid [z_1,\ z_2]^T \in \mathcal{Z}\}, \tag{15c}$$
$$\overline{y}_2 = \max\{f_1(z_1) - f_2(z_2) \mid [z_1,\ z_2]^T \in \mathcal{Z}\}, \tag{15d}$$

which can be computed by solving four NLP problems. Finally, since all expressions are now functions of a single variable, the PWA approximations $\tilde{f}_1(z_1) \approx f_1(z_1)$, $\tilde{f}_2(z_2) \approx f_2(z_2)$, $\tilde{f}_{y_1}(y_1) \approx f_{y_1}(y_1)$, and $\tilde{f}_{y_2}(y_2) \approx f_{y_2}(y_2)$ can be computed by solving the NLP (10). The overall optimal PWA approximation $\tilde{f}(z_1, z_2) \approx f(z_1, z_2)$ then becomes

$$\tilde{f}(z_1, z_2) = {}^1\!/\!_4\Big(\tilde{f}_{y_1}\big(\tilde{f}_1(z_1)+\tilde{f}_2(z_2)\big)-\tilde{f}_{y_2}\big(\tilde{f}_1(z_1)-\tilde{f}_2(z_2)\big)\Big). \tag{16}$$

The evaluation procedure is similar as above. I.e., given the arguments z_1 and z_2, one first evaluates $\tilde{z}_1 = \tilde{f}_1(z_1)$ and $\tilde{z}_2 = \tilde{f}_2(z_2)$. Subsequently, one evaluates $\tilde{y}_1 = \tilde{f}_{y_1}(\cdot)$ with the argument $\tilde{z}_1 + \tilde{z}_2$, then $\tilde{y}_2 = \tilde{f}_{y_2}(\cdot)$ at the point $\tilde{z}_1 - \tilde{z}_2$. Finally, $\tilde{f}(z_1, z_2) = {}^1\!/\!_4(\tilde{y}_1 - \tilde{y}_2)$.

4.3 Multivariable Nonseparable Functions

When the nonlinear function $f : \mathbb{R}^{n_z} \to \mathbb{R}$ to be approximated does not satisfy Assumption 4.1, we propose to proceed as follows. As a rather general setup, consider that

$$f(z) = f_{\mathrm{out},1}(f_{\mathrm{out},2}(f_{\mathrm{out},3}(\cdots(f_{\mathrm{in}}(z))))) \tag{17}$$

with the inner function $f_{\mathrm{in}} : \mathbb{R}^{n_z} \to \mathbb{R}$ satisfying Assumption 4.1 and arbitrary outer functions $f_{\mathrm{out},i} : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, m-1$. This relation can be further generalized to include sums and/or products of functions.

As a specific example, consider

$$f(z_1, z_2) = \exp(z_1 z_2), \tag{18}$$

where $f_{\mathrm{in}}(z_1, z_2) = z_1 z_2$ and $f_{\mathrm{out}}(w) = \exp(w)$. To derive an optimal PWA approximation $\tilde{f}$ of (18) , we introduce the substitution $w = f_{\mathrm{in}}(z_1, z_2)$. Since f_{in} satisfies Assumption 4.1, the procedure of Section 4.2 can be applied to find its optimal PWA approximation $\tilde{w} = \tilde{f}_{\mathrm{in}}(z_1, z_2) \approx z_1 z_2$. Define two new variables $y_1 = (z_1+z_2)$ and $y_2 = (z_1 - z_2)$. Then ${}^1\!/\!_4(y_1^2 - y_2^2) = z_1 z_2$ trivially holds. Subsequently we can solve the NLP (10) to obtain optimal PWA approximations $\tilde{f}_{y_1}(y) \approx y^2$ on domain $[\underline{y}_1, \overline{y}_1]$ and $\tilde{f}_{y_2}(y) \approx y^2$ on domain $[\underline{y}_2, \overline{y}_2]$. We remark that although both functions to be approximated are the same (y^2), their respective domains will be different and are given by (15). Their PWA approximations will therefore differ as well. Next we derive a PWA approximation of $\tilde{f}_{\mathrm{out}}(w) \approx \exp(w)$ again by solving (10). Value of the overall PWA approximation $\tilde{f}(z_1, z_2) \approx \exp(z_1 z_2)$ at a particular point $(z_1,\ z_2)$ can then be obtained by evaluating the corresponding 1D approximations in the following order:

1. $\tilde{y}_1 = \tilde{f}_{y_1}(z_1 + z_2)$
2. $\tilde{y}_2 = \tilde{f}_{y_2}(z_1 - z_2)$
3. $\tilde{w} = {}^1\!/\!_4(\tilde{y}_1 - \tilde{y}_2)$
4. $\tilde{f}(z_1, z_2) = \tilde{f}_{\mathrm{out}}(\tilde{w})$

Such an substitution approach can be generalized to derive optimal PWA approximations of general nonlinear functions in the form of (17) by the following procedure:

1. Obtain optimal PWA approximation of the inner function $f_m(z)$ using the procedure in Section 4.2.
2. Define new variables w_i and approximate the 1D functions $f_i(w_i)$, $i = m-1, \ldots, 1$, by solving (10).

If the multivariable inner function $f_{\mathrm{in}} : \mathbb{R}^{n_z} \to \mathbb{R}$ with domain $\mathcal{Z}$ consists of more than two terms, its PWA approximation can be performed in an inductive manner. Consider $f_{\mathrm{in}}(z_1, z_2, z_3) = f_1(z_1)f_2(z_2)f_3(z_3)$. First, approximate the product $f_1(z_1)f_2(z_2)$ by a PWA function of the form of (16), which requires four PWA approximations $\tilde{f}_1(\cdot) \approx f_1(\cdot)$, $\tilde{f}_2(\cdot) \approx f_2(\cdot)$, $\tilde{f}_{y_1}(\cdot) \approx y_1^2$, $\tilde{f}_{y_2}(\cdot) \approx y_2^2$, where y_1 and y_2 are as in (14). Let $f_a(z_1, z_2) := f_1(z_1)f_2(z_2)$. Then $f(z_1, z_2, z_3) = f_a(z_1, z_2)f_3(z_3)$, which can again be approximated as a product of two functions. Specifically, define

$$y_3 = f_a(\cdot) + f_3(z_3), \quad y_4 = f_a(\cdot) - f_3(z_3), \tag{19}$$

and hence $f_a(z_1, z_2)f_3(z_3) = {}^1\!/\!_4(y_3^2 - y_4^2)$. The domains over which y_3^2 and y_4^2 need to be approximated are, respectively, $[\underline{y}_3,\ \overline{y}_3]$ and $[\underline{y}_4,\ \overline{y}_4]$ with

$$\underline{y}_3 = \min\{f_1(z_1)f_2(z_2) + f_3(z_3) \mid z \in \mathcal{Z}\}, \tag{20a}$$
$$\overline{y}_3 = \max\{f_1(z_1)f_2(z_2) + f_3(z_3) \mid z \in \mathcal{Z}\}, \tag{20b}$$
$$\underline{y}_4 = \min\{f_1(z_1)f_2(z_2) - f_3(z_3) \mid z \in \mathcal{Z}\}, \tag{20c}$$
$$\overline{y}_4 = \max\{f_1(z_1)f_2(z_2) - f_3(z_3) \mid z \in \mathcal{Z}\}, \tag{20d}$$

and $z = [z_1,\ z_2,\ z_3]^T$. Subsequently, three additional PWA approximations

$$\tilde{f}_{y_3}(y_3) \approx y_3^2,\ \tilde{f}_{y_4}(y_4) \approx y_4^2,\ \tilde{f}_3(z_3) \approx f_3(z_3)$$

need to be computed over the corresponding domains. The aggregated optimal PWA approximation $\tilde{f}(z_1, z_2, z_3) \approx f(z_1)f(z_2)f(z_3)$ consists of 7 individual approximations and is given by

$$\tilde{f}_{\text{in}}(\cdot) = {}^1\!/_4\Big(\underbrace{\tilde{f}_{y_3}\big(\hat{f}_a + \tilde{f}_3(z_3)\big)}_{\hat{y}_3} - \underbrace{\tilde{f}_{y_4}\big(\hat{f}_a - \tilde{f}_4(z_3)\big)}_{\hat{y}_4}\Big). \quad (21)$$

Here, $\hat{f}_a$ is the function value of $\tilde{f}_a(z_1, z_2) \approx f_1(z_1)f_2(z_2)$ at z_1 and z_2, where $\tilde{f}_a(\cdot)$ is obtained from (16), i.e.:

$$\hat{f}_a = {}^1\!/_4\Big(\underbrace{\tilde{f}_{y_1}\big(\tilde{f}_1(z_1) + \tilde{f}_2(z_2)\big)}_{\hat{y}_1} - \underbrace{\tilde{f}_{y_2}\big(\tilde{f}_1(z_1) - \tilde{f}_2(z_2)\big)}_{\hat{y}_2}\Big). \quad (22)$$

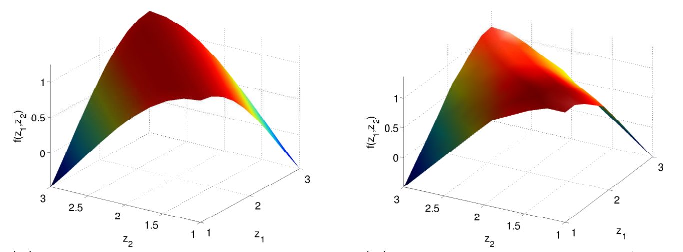
The overall PWA approximation $\tilde{f}_{\text{in}}(z_1, z_2, z_3)$ can then be evaluated, for any z_1, z_2, $z_3 \in \mathcal{Z}$, by computing the function values of the respective approximations.

Such an inductive procedure can be repeated *ad-infinitum* to derive PWA approximations of any multivariable inner function. In general, the PWA approximation will consists of $2p + n_z + m - 1$ individual PWA functions, where n_z is the number of variables, m is the number of functions in (17) and p is the number of products between individual subfunctions $f_j(z_j)$ in the inner function f_{in}. As an example, for $f_{\text{in}}(z) := \alpha_1 f_1(z_1)f_2(z_2)f_4(z_4) + \alpha_2 f_3(z_3)f_5(z_5)$ we have $p = 3$. We remark that inclusion of scalar multipliers α_j into the PWA description of the form (21)–(22) is straightforward and only requires linear scaling of the corresponding terms.

Example 4.2. Consider a set of input-ouput data in $\mathbb{R}^2$. To fit these data with a PWA function, we have first applied the procedure of Section 3 to obtain an optimal fit by the function$f(z) = \sum_{i=1}^{3} \alpha_i f_i(z)$ which consists of basis functions $f_1 = 1$, $f_2 = \sin(z_1 z_2)$ and $f_3 = \cos(z_1 - z_2)$. By minimizing (3) we have obtained

$$f(z_1, z_2) = 0.02 + 0.08\sin(z_1 z_2) + 1.2\cos(z_1 - z_2), \quad (23)$$

shown in Figure 2(a). To derive an optimal PWA approximation of f in (23) we have applied the aforementioned procedure to approximate $\sin(z_1 z_2)$ by first approximating $z_1 z_2$ by a PWA function $\tilde{f}_1(z_1, z_2)$ and $\sin(w)$ by $\tilde{f}_2(w)$. Approximation of $\cos(z_1 - z_2)$ was performed in a similar manner. The resulting PWA approximation of (23), consisting of 15 regions, is depicted in Figure 2(b).



(a) Optimal fit with a trigonometric polynomial basis.

(b) Graph of optimal PWA approximation $\tilde{f}(z_1, z_2)$

Fig. 2. Two-dimensional fit from Example 4.2.

Remark 4.2. Approximation methods based on approximation of the domain, e.g. Delaunay triangularization is based on the fact, that each point will be associated with one triangle. So the number of regions as well as the parameters of each PWA approximation grows linearly with the number of points, while our method does not have to take into account such a restriction. In n-dimensional case, these methods could have some advantages related to the number of regions, when the cardinality of the set, representing the input-output measurements is small. Nevertheless, obtaining the parameters of the hyperplanes, representing the PWA approximation in higher dimensions is trickier, since our method is based on solving a series of one-dimensional approximation, hence the whole problem is reduced to seeking the parameters of lines.

5. CASE STUDY

Consider a continuous stirred tank reactor (CSTR) where the reaction $A \rightarrow B$ takes place. The source compound is pumped into the reactor at a constant inflow with a constant concentration. The chemical reaction is exothermic and a coolant liquid is therefore pumped into the reactor's jacket to prevent overheating. The input temperature of the coolant is constant, while its flow rate q_c can be manipulated and is considered an exogenous input. Concentration of the reactant c_A inside of the reactor, temperature of the reactor mixture ϑ, and temperature of the cooling liquid in the jacket ϑ_c are the state variables of the CSTR. The normalized material and energy balances of such a reactor are then given by

$$\begin{aligned} \dot{c}_A &= \alpha_1 - \alpha_2 c_A - \gamma(c_A, \vartheta), \\ \dot{\vartheta} &= \alpha_4 - \alpha_5\gamma(c_A, \vartheta) + \alpha_6\vartheta + \alpha_7\vartheta_c, \\ \dot{\vartheta}_c &= \alpha_8 q_c + \alpha_9(\vartheta - \vartheta_c) - \alpha_{10}\vartheta_c q_c, \end{aligned} \quad (24)$$

with $\gamma(c_a, \vartheta) = \alpha_3 c_A e^{-\beta/\vartheta}$. Values of all parameters, except of α_3 and β, are known from available chemical-engineering sources, but coefficients α_3 and β can only be determined from experimental measurements.

To find an optimal PWA approximation of (24) we have first applied the procedure of Section 3 to obtain an analytic expression of the nonlinear function γ, along with concrete numerical parameters. The obtained optimal fit was in the form of

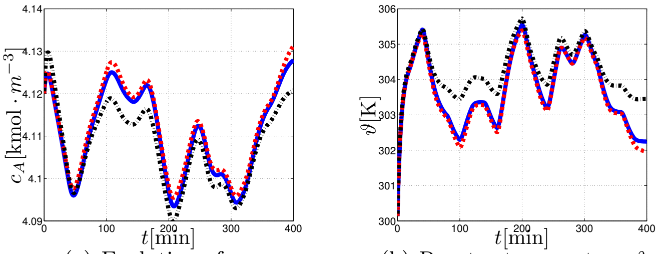
$$\gamma(c_A, \vartheta) = 4.7772 \cdot 10^{13} c_A e^{-11500/\vartheta}. \quad (25)$$

Next, we have applied the procedure described in Section 4 to obtain an optimal PWA approximation of right-hand-sides of the nonlinear model (24)-(25). The model contains two nonlinear terms to be approximated: the product between jacket temperature and coolant inflow ($\vartheta_c q_c$) and the nonlinear reaction rate $\gamma(c_A, \vartheta)$ in (25). The first nonlinearity satisfies Assumption 4.1 and its PWA approximation can therefore be obtained as described in Section 4.2. Approximation of γ is more involved, since it does not satisfy Assumption 4.1. Therefore we have to apply the procedure of Section 4.3 by rewriting (25) as $\gamma(c_A, \vartheta) = \alpha_3 c_A f_{\text{out}}(f_{\text{in}}(\vartheta))$, where $f_{\text{out}}(w) = e^w$ and $w = f_{\text{in}}(\vartheta) = -\beta/\vartheta$. Here, the inner function $f_{\text{in}}(\vartheta)$ is a function of single variable, therefore its PWA approximation $\tilde{f}_{\text{in}}$ could be obtained per solving (10). PWA approximation of the outer function $\tilde{f}_{\text{out}}$ can be obtained in the same manner. Finally, the product $c_A f_{\text{out}}(w)$ meets requirements of Assumption 4.1 and therefore its optimal PWA approximation is obtained from (16).

The overall approximation of the CSTR model (24) is then given by

$$
\begin{aligned}
\dot{c}_A &\approx \alpha_1 - \alpha_2 c_A - \alpha_3 \tilde{\gamma}(c_A, \vartheta), \\
\dot{\vartheta} &\approx \alpha_4 - \alpha_5 \tilde{\gamma}(c_A, \vartheta) + \alpha_6 \vartheta + \alpha_7 \vartheta_c, \\
\dot{\vartheta}_c &\approx \alpha_8 q_c + \alpha_9 (\vartheta - \vartheta_c) - \alpha_{10} \tilde{f}(\vartheta_c, q_c),
\end{aligned}
\tag{26}
$$

where $\tilde{\gamma}$ is the approximation of γ and $\tilde{f}(\vartheta_c, q_c) \approx \vartheta_c q_c$.



(a) Evolution of c_A.

(b) Reactor temperature ϑ.

Fig. 3. Simulation results for the CSTR: nonlinear model (24) (red line), PWA approximation (26) (red dashed line), linear approximation (black dotted line).

To assess approximation accuracy, we have investigated the open-loop evolution of the original nonlinear model (24) and compared it to the behavior of its PWA approximation (26) with 10 regions. Time evolution of two state variables are shown in Figure 3. To better illustrate advantages of the PWA approximation, the simulation scenario also shows evolution of linearized version of (24) around the nominal steady state As can be seen from the results, the PWA approximation clearly outperforms the model based on a single linearization. Important to notice is that the PWA model consists of 10 local linear models.

6. CONCLUSION

In this paper we have shown how to derive PWA approximations of arbitrary nonlinear relations from measured data. The procedure consists of two steps. In the first part the data are first fit with a nonlinear function by minimizing the fitting error and, alternatively, also minimizing complexity of the fitting function. In the second part we have shown how to employ nonlinear optimization to derive optimal PWA approximations of arbitrary nonlinear functions. Specifically, we have illustrated that approximation of multivariable functions boils down to a series of one-dimensional approximations with a favorable complexity. Moreover, we have reported how to approximate functions which are not separable. Procedures and algorithms reported in this paper are available in our AUTOPROX toolbox, which is available for free download from `http://www.kirp.chtf.stuba.sk/~sw/`. The toolbox provides an easy-to-use interface to derivation of optimal PWA approximations and is also capable to exporting the resulting models into the HYSDEL language.

ACKNOWLEDGMENT

The authors are pleased to acknowledge the financial support of the Scientific Grant Agency of the Slovak Republic under grants 1/0095/11 and 1/0973/12.

REFERENCES

Adjiman, C.S., Androulakis, I.P., Maranas, C.D., and Floudas, C.A. (1996). A global optimization method, αBB for process design. *Global Optimization of Dynamic Systems*, 20, 419–424.

Bemporad, A. and Morari, M. (1999). Control of systems integrating logic, dynamics, and constraints. *Automatica*, 35(3), 407–427.

Boyd, S. and Vandenberghe, L. (2004). *Convex optimization*. Cambridge University Press.

Branicky, M.S. (1995). *Studies in Hybrid Systems: Modeling, Analysis, and Control*. Ph.D. thesis, Massachusetts Institute of Technology.

Chachuat, B., Singer, A.B., and Barton, P.I. (2006). Global methods for dynamic optimization and mixed-integer dynamic optimization. *Ind. Eng. Chem. Res.*, 45(25), 8373–8392.

Corona, D. and De Schutter, B. (2008). Adaptive cruise control for a SMART car: A comparison benchmark for MPC-PWA control methods. *IEEE Transactions on Control Systems Technology*, 16(2), 365–372.

Ferrari-Trecate, G., Muselli, M., Liberati, D., and Morari, M. (2001). Identification of Piecewise Affine and Hybrid Systems. In *American Control Conference*, 3521–3526. Arlington (VA), USA.

Gegúndez, M., Aroba, J., and Bravo, J. (2008). Identification of piecewise affine systems by means of fuzzy clustering and competitive learning. *Engineering Applications of Artificial Intelligence*, 21(8), 1321–1329.

ILOG, Inc. (2003). *CPLEX User Manual*. Gentilly Cedex, France.

Kozák, S. and Števek, J. (2011). Improved piecewise linear approximation of nonlinear functions in hybrid control. In *Preprints of the 18th IFAC World Congress Milano (Italy) August 28 - September 2, 2011*, volume 18, 14982–14987.

Kvasnica, M., Szũcs, A., and Fikar, M. (2011). Automatic derivation of optimal piecewise affine approximations of nonlinear systems. In *Preprints of the 18th IFAC World Congress Milano (Italy) August 28 - September 2, 2011*, 8675–8680.

Löfberg, J. (2004). YALMIP. Available from `http://users.isy.liu.se/johanl/yalmip/`.

Ohlsson, H., Ljung, L., and Ohlsson, H. (2010). Identification of piecewise affine systems using sum-of-norms regularization. *Regularization for Sparseness and Smoothness*, 109.

Paoletti, S., Juloski, A., Ferrari-Trecate, G., and Vidal, R. (2007). Identification of hybrid systems: A tutorial. *European Journal of Control*, 13(2-3), 242–260.

Papafotiou, G., Geyer, T., and Morari, M. (2007). A hybrid model predictive control approach to the direct torque control problem of induction motors. *International Journal of Robust & Nonlinear Control*, 17, 1572–1589.

Papamichail, I. and Adjiman, C.S. (2004). Global optimization of dynamic systems. *Computers and Chemical Engineering*, 28, 403–415.

Roll, J., Bemporad, A., and Ljung, L. (2004). Identification of piecewise affine systems via mixed-integer programming. *Automatica*, 40(1), 37–50.

Sontag, E.D. (1981). Nonlinear regulation: The piecewise linear approach. *IEEE Transactions on Automatic Control*, 26(2), 346–358.

Torrisi, F. and Bemporad, A. (2004). HYSDEL — A tool for generating computational hybrid models for analysis and synthesis problems. *IEEE Transactions on Control Systems Technology*, 12, 235–249.

Williams, H. (1993). *Model Building in Mathematical Programming*. John Wiley & Sons, Third Edition.